

Experiences in Re-engineering the Approach to Editing and Imputing Canadian Imports Data

F. Laflamme, C. Barrett, W. Johnson and L. Ramsey¹
Statistics Canada

Abstract

The International Trade Division of Statistics Canada has recently re-engineered the Edit and Imputation strategy of the Imports Program. Essentially based on the micro edit approach, the old strategy did not meet the current needs regarding the efficient use of resources and the flexibility of adaptation to new requirements, necessary for the management of so vast a statistical program. The new approach is inspired by the macro edit techniques without forsaking those of micro editing. Indeed, these two different approaches will be used in a complementary way within the proposed system. This paper will describe the new modular system to edit the Canadian import data and its main characteristics including some of the elements that have promoted the choice of this new orientation more directed towards macro editing.

Key words: Macro edit, micro edit, selective edit, aggregates

1. Introduction

The edit strategy of International Trade Division (ITD) of Statistics Canada was developed during the seventies. The approach followed the edit ideas prevalent during this period: examine each and every record. By the late eighties, budget reductions had forced ITD to move away from this strategy; only records above a certain threshold were manually reviewed. This approach was used to reduce the number of records selected for review to react to successive budgetary constraints. But the original idea remained the same: all selected records were edited completely.

There are several reasons why the International Trade Division concluded that its approach to editing imports data needed to be re-evaluated. The ever-increasing volume of administrative data as a result of the growth in trade activity and a greater number of transactions due to "just-in-time" production; the need to use limited operational resources more efficiently; and the requirement for a flexible system are some of these reasons. The Memorandum of Understanding on exchange of Imports Data between Statistics Canada and the United States Bureau of Census (USBC) and the re-engineering at Revenue Canada Customs, Excise and Taxation, the main source of imports data, also contributed to the need for re-engineering the ITD edit strategy.

The objectives of the re-engineering exercise must not be based only on the desire to save money, but must maintain improvements to data quality and ensure timeliness of outputs. The challenge is to find the right balance in terms of quality, the need to avoid delays of publication, the operational constraints, and the cost. Granquist and Kovar (1994) give a very good discussion of these concerns.

The proposed system should target more efficiently the records subjected to manual review. The approach should also be flexible enough to easily integrate new requirements that may affect the ITD program. The workload, in terms of number of records reviewed, should be monitored and controlled over time to ensure that ITD consistently reaches its publication deadlines.

An analysis of the structure of the Canadian import data and the impact of the current edit system led to the decision to base the new edit system on the macro edit approach; elements of the micro edit approach will continue to be important. Chinnappa et al. (1990) states: "In a general sense, editing includes all activities that involve checking for the integrity, consistency and coherence of the data at aggregate levels (macro editing) or at record level (micro editing)". Pursey (1994) and Laflamme (1995) provide discussions of methods of editing for the ITD imports data.

¹ F. Laflamme, Business Survey Methods Division, 11th Floor, R.H. Coats Building; C. Barrett, W. Johnson and L. Ramsey, International Trade Division, 9th Floor, Jean Talon Building, Tunney's Pasture, Ottawa, Ontario K1A 0T6.

2. Current process

The type of data managed in the Canadian imports program is composed of two categories: qualitative, such as type of product, unit of measure, country of origin, mode of transport and importer number; and quantitative as in the value and quantity of the product. The product code is based on the Harmonized System (HS) that provides a structured and exhaustive list of goods to classify. Its structure is hierarchical and the level of detail increases with the level of classification. For example, HS2 (first 2 digits) might represent "Live Animals" while HS10 would more specifically refer to "Pure-bred Breeding Animals, Dairy". As of January 1996, the number of possible codes for the HS2, HS4, HS6, HS8 and HS10 are 98, 1 268, 5 156, 8 492, and 17 165 respectively. ITD also creates nine "specialty" groupings by aggregating one or more HS2(chapter) levels.

International Trade Division receives administrative data related to Imports on a weekly basis from Customs. These data are collected at all Canadian ports of entry through Custom's "B3" document and directly by importers through Customs' Cadex system. Data are also received from other sources such as the Postal Imports Control System (PICS) and the National Energy Board (NEB).

The B3 data received on tape are first processed through a mainframe system known as the "preprocessor". This system reformats and "cleans" the data as well as removes shipments considered "non-trade" and aggregates small value shipments. After the preprocessor, the data are loaded on to the Concurrent Mini-computer where edits are applied in batch mode.

The first set of edits verify specific fields by matching to validity tables. Edits which verify specific combinations of fields are applied next. The final edit applied is unit value. The unit price per record is validated to determine whether it falls within the established range by HS code.

There are three reasons why a record can be chosen for manual review. A record is possibly non-trade or requires monitoring; an edit has failed and the value of a record is above the level for automatic system imputation or there is no edit failure but a record is above a specified value or quantity. The records requiring review are then presented line-by-line by specialty grouping. Paper documents (B3 and invoice) are retrieved for all of these transactions and provided to the reviewer. Results of the edits are shown on the screen to assist in the on-line review. After corrections are made the record is edited again and if considered "clean", stored on the base. Flags are used to record the outcome of the batch edits, imputations and on-line corrections.

3. Overview of the new system

The new strategy will use the macro edit approach extensively. However, it will continue to use a micro edit approach to ensure that the data from micro records sum to the macro estimates. It will efficiently use edit resources by targeting more directly the records requiring manual review. These records represent a large proportion of the overall cost associated with data editing. One of the challenges was to use jointly the micro and macro approaches. The new edit system flow will integrate in an innovative and efficient way these two distinct but complementary approaches (Laflamme et al., 1996).

The edit system uses a series of modules that give flexibility in integrating edit requirements and also controlling the number of records manually reviewed. Each module has a specific function in the overall system and serves to target in different ways the records to be manually reviewed. High impact records, records with micro edit failures that belong to aggregates with a high potential error, comparison of current aggregates to historical aggregates, and other requirements will be successively handled in the edit system. Tools for both efficient monitoring of the edit process as well as data analysis have been developed in parallel to the system ensuring that data changes make sense in general (Laflamme et al., 1995). The edit system flow is presented on page 5.

4. Key characteristics of the system

The fact that the system will efficiently target the records subjected to review is one key characteristic. The facilities in terms of the integration of the new requirements in the different modules, the ability to handle some specific demands, and the use of the acquired subject matter expertise in a convenient way are definitively three other positive characteristics. The

general flexibility of the overall system is probably among the main characteristics and advantages of the strategy. The capacity to modify, update or remove most of the parameters of any module provides a great flexibility regarding the workload distribution among the edit modules. The possibility to control the number of records manually reviewed in each module and between each of them gives to the ITD management an efficient tool to regulate the allocation of the editing resources. It also provides the ability to adapt the system easily and quickly to any future budgetary cost reduction. The uniformity of the workload distribution over time could also be achieved by the control of the number of records selected for review in each module.

Another key characteristic is the “system memory”. Since the approach is sequential, the system will consider the work done in the previous modules. In this way, we will again target, more efficiently, the records that have the high potential impact on the data at any given point in time and we will avoid as much as possible, the overlap or duplication in the edit process.

The proposed approach systematically takes care of the three complementary dimensions of the program: high impact records, high potential error aggregates and historical data comparison. It also uses several edit approaches in the same system: selective edit, micro edit, macro edit and a combination of micro and macro edits in the pre-macro edit module.

The modularity of the system is one of the characteristics that provides many significant advantages. First, it promotes the development of new modules and the improvement of existing ones. Second, it reduces the risk of introducing error because only one module should be modified or updated at a time. Finally, the modularity favours both the comparison of different edit strategies in terms of the specific parameters used in a given module and the editing resources allocation e.g. the number of records allowed to be reviewed in each module of the system.

The event and audit logs, which record the results of the edit, imputation and review processes, is another important feature of the system. The logs will make available to the analyst all the information required to determine whether the edits are working effectively and if the imputations and manual revisions are acceptable.

Finally, the overall reduction of the number of records subject to manual review combined with the new optional approach of the document retrieval procedure should considerably reduce the operating cost of the edit and imputation process.

5. Factors that support the new orientation

The volume and the skewed distribution of data, the contribution of large records-in-error to the total error, the contribution of the records-in-error to the aggregate to which they belong and the concentration of the error in the aggregate are some of the factors that support the new orientation.

The ITD processes about 1.3 million records on a monthly basis for a total imports value of 16 billion dollars. The distribution of the data is very skewed with 1.7 % of the records representing approximately 47% of the imports value for a given month. Therefore, the error in this type of record has a very large impact on the aggregate totals. We have evaluated that the records of \$500,000 and over that contain error did contribute to more than 50% of the value that pass from an aggregate to another aggregate in the edit process. This measure, called relative flow, is based on a transition matrix that shows the transfer of value for each aggregate during the editing process for a given level of aggregation. The use of this method of evaluation assumes that only the attributes are changed, for example, the product or the country of origin, not the value of the record. In fact, the value of about 300 records is modified each month representing approximately 0.2% of the total imports value, which is negligible.

A concentration of errors in some domains of activities is also observed in the data. Descriptive analysis shows a concentration of errors, for example in the machinery and transportation sectors.

The macro edit approach will be useful because it is possible to identify the groups of records that have a large impact on the aggregates and also some groups commonly in error.

6. Description of the edit and imputation system

Figure 1 outlines the new edit and imputation process. This process will begin after pre-processing of the Custom's

tape that is received weekly from Customs. The first function of the system is the validation of the data, and continues through each of the modules exiting from the edit and imputation system after the USBC edits have been applied. The data are now ready for the post edit and imputation process, principally the creation of the US files for transmission to the USBC and the creation of detailed, major and summary group files for Canadian Trade Statistics.

There are four modules in the edit and imputation process that will result in the identification of records for review by the data quality analysis staff: selective edits, pre-macro (objective) edits, pre-macro (subjective) edits and the macro edits. There is one other group of records to be identified for review in the pre-process function. These are records requiring a decision on whether the data should be included in the trade statistics or be considered a non-trade item. An example of a non-trade commodity would be an item that is temporarily imported for a trade show and will be returned to the originating country after a short period of time.

The outcome of the edit, imputation and manual review processes will be tracked and recorded in “log” tables. These tables will be used to monitor the efficiency of the edit process and to identify possible data quality problems.

Each of the editing modules are described individually.

6.1 Validation/Imputation edit module

Selected data fields will be checked to verify that the values are of the correct format, numeric, alpha, blank etc. This stage will also confirm that the value entered in any field is valid, where a table of valid codes is available. If any of the fields fail the validation edit, their value will be imputed by the system at this stage. This is the first step of the edit and imputation process and it must guarantee that the data can be aggregated by any attribute of any field.

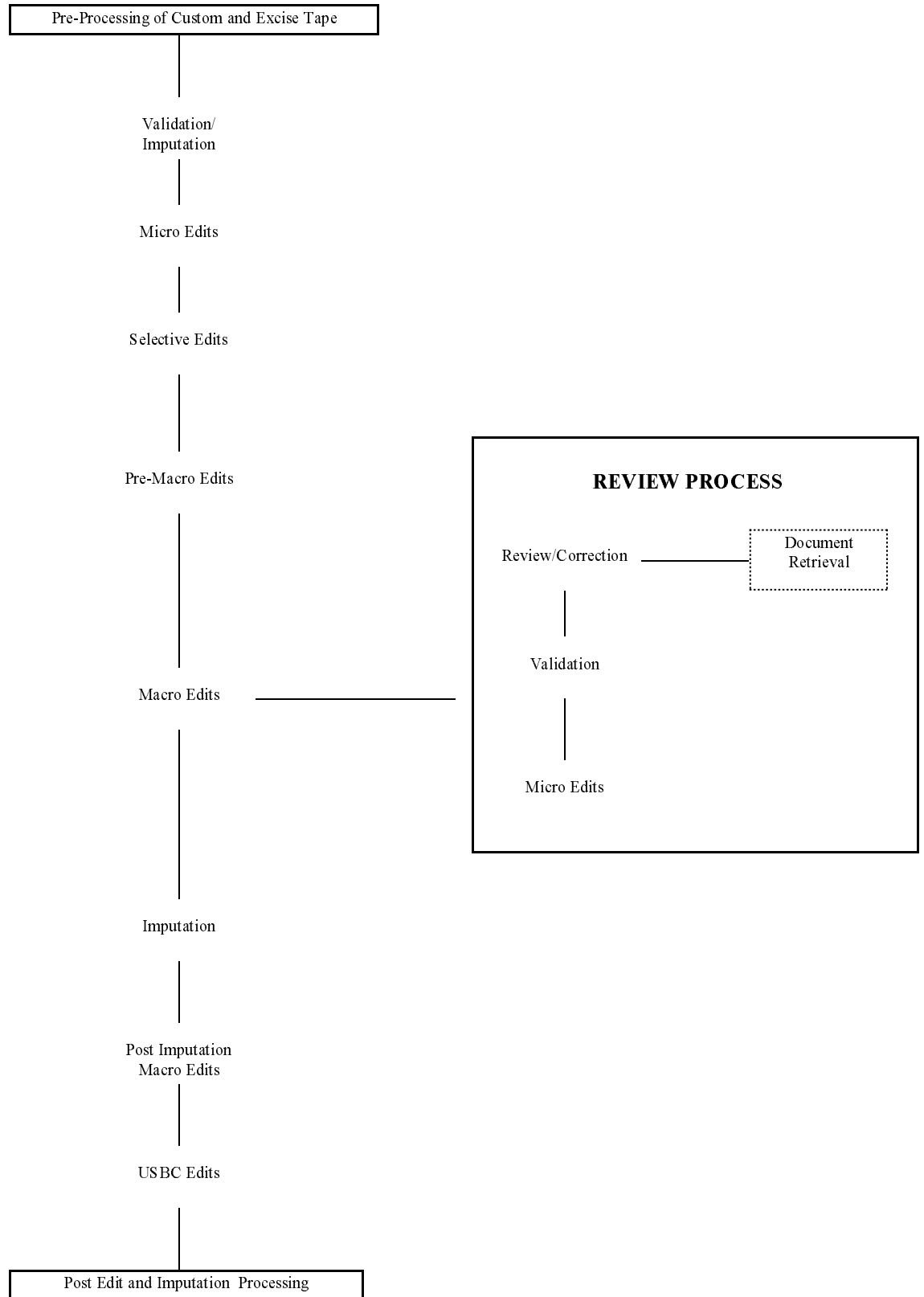
6.2 Micro edit module

The purpose of the micro level edits is to determine if the information reported is consistent within the record. Identification of potential errors is recorded at this stage, although no manual or system intervention is employed.

There will be seven micro edits performed on the records to determine if the information is consistent. The results of the micro edits will be used in the pre-macro module to identify important aggregates in terms of value that contain the highest rate of potential error for each of the pre-selected attributes. During any review of the records (resulting from selective, pre-macro and macro edits), the identifiers that have been set during this stage will serve as a tool to assist in pinpointing problem fields.

The micro edits consist of HS10/Country, HS10/Importer, HS10/Unit of Measure, Mode of Transport(Marine)/Carrier Code at Importation (Special), Mode of Transport(Air)/Carrier Code at Importation, Mode of Transport(Marine)/Carrier Code at Importation (Regular), and Mode of Transport(Mail)/Entry Type. For example, the HS10/Country edit verifies that the HS10 and Country combination is acceptable. The table (HS10/Country Exclusion) contains a list of countries from which the HS10 cannot be imported. If the combination is found on this table, the edit has failed. All records not reviewed in subsequent modules and contain micro edit failures will be imputed prior to closing the processing cycle.

Figure 1. PRODUCTION PROCESS EDIT AND IMPUTATION IMPORTS



6.3 Selective edit module

The selective edit is applied to the data on a weekly basis following the micro edits. The purpose of the selective edit is to identify for review, the most important records in terms of value and quantity. A large portion (50%) of the relative flow comes from a small number of high value records. The selective edit allows these records to be checked early in the process assuring a certain level of data quality. Reviewing these records at this time will account for a majority of the change and is independent of any edit result.

An overall number of records which will be selected for review is determined at the start of the cycle. This number is allocated over each chapter and further distributed over each tape within the cycle, selecting the highest value and highest quantity records within each. The overall number of records to be selected is not a fixed number and can be adjusted depending on workload constraints.

The distribution of the overall number of records for the cycle is determined using the square root of the annual chapter values. A minimum number of 16 records will be assigned to each chapter.

Approximate total : 3600 units (N) Square Root Allocation between chapters: $X_i = (\text{Chapter}_i \text{ value})^{0.5} / \sum_i (\text{Chapter}_i \text{ value})^{0.5} \quad i=1,2,\dots,99$				
	Chapter Allocation	Tape Allocation	Value Allocation	Quantity Allocation
Chapter _i	$Z_i = \max(16, X_i N)$	$T_k = \text{Int}((Z_i - \sum_{k=1} T_{(k-1)}) / (4-k+1))$ $k=1,2,3,4$	$V_k = \text{Int}(0.75 T_k)$	$Q_k = 1 - V_k$

The selective edit further targets high value as well as high quantity records. In the above table, approximately 75% of the review records will be selected for high value and the remaining for high quantity. The parameters in the table can be changed to adjust the total records for review from the selective edit, the distribution among chapters, including a minimum number and the split between high value and high quantity records.

The records from the first tape will be sorted by descending value and the record allocations applied. All of the records will then be sorted again by descending quantity. The process for the 2nd to 4th tapes is the same, with one exception. The records from these tapes will be concatenated with all records not selected from previous tapes, ensuring that the most important high value and high quantity records are always selected within each cycle. The selected records will be presented for review by specialty (chapter groupings).

6.4 Pre-Macro (Objective) edit module

The aim of the pre-macro (objective) edit is to identify aggregates that contain a high degree of potential errors. The potential errors were identified at the record level in the micro edit module. The potential micro errors relate to combinations of fields; specifically, HS/Country, HS/Importer, HS/UOM, MOT/Carrier Code, MOT/Entry Type and HS/Unit Value, with the HS/Unit Value edit further defined by Country and Importer.

All aggregates at the HS10 level combined with select edit field variables (i.e. Country, Importer etc.) that have a potential error will be calculated and divided by the total value of the HS10 to produce a ratio. This ratio will be weighted by a function of the total value of the HS10. The weighted ratios will then be ranked in descending order. The highest ranked 50 aggregates for each attribute will be selected. From each of these 50 aggregates, the three largest records in terms of value that contain potential error will be selected for review.

There will be seven attributes times a maximum of 3 records times fifty aggregates resulting in a maximum of 1,050 records for review per week.

Example:

Consider that HS10A and HS10B contain five errors each. The total value of the five records with potential HS/Country errors belonging to HS10A is \$3,000. The total value of the five records belonging to HS10B with potential HS/Country errors is \$5,000. Along with the five HS/Country error records for HS10A are 10 records with a value of \$7,000. Thus the total value of all HS10A record equals \$10,000 (7,000 + 3,000). For HS10B there are 10 more records worth \$20,000 for a total value for this HS of \$25,000 (20,000 + 5,000).

The following table summarizes the information for the HS/Country pre-macro edit:

HS	Ratio	Weighted Ratio
10A	$3000/10000 = .30$	<input type="text"/> = 30.0
10B	$5000/25000 = .20$	<input type="text"/> = 31.6

The weighted ratio associated with HS10B is greater than the weighted ratio associated with HS10A and would be ranked higher, although the ratio of potential error is greater for HS10B. All HS10 aggregates that have any potential error for the HS/Country micro edit would have the ratio calculated. The highest ranked fifty aggregates would be identified as suspicious aggregates. A maximum of three records from each aggregate will be selected for review. Each of the other attributes would be treated in the exact same manner.

The pre-determined number of records for review from the pre-macro (objective), can be adjusted at the total level or a change in the distribution among edit attributes can be modified.

6.5 Pre-macro (Subjective) edit module

There are situations that are known to the data quality analysis staff of data fields reported in a transaction are suspect. This information is stored in the monitoring table for the pre-macro (subjective) edit. The majority of monitoring criteria exists to verify the HS Classification, although any field can be a candidate to identify records for review. The responsible staff have particular knowledge of suspicious data and use the monitoring tool to select these records for review and perform the appropriate correction. A concrete example involves a specific importer, from a particular state who at times incorrectly classifies a commodity. All records exactly matching this criteria will be selected for review. The number of records for review from the pre-macro subjective edit is totally flexible, in that it is the responsibility of the subject matter specialist to monitor and ensure the relevance of the criteria used.

6.6 Macro edit module

The macro edit module targets records that belong to the most suspicious aggregates, when compared to the closest value during the previous twelve months for these aggregates.

The specific attributes and levels of aggregation for the macro module are: HS10, HS2/Country, HS2/Importer and HS2/Mode of Transport. Each of these four attributes will be compared to their previous history in terms of total monthly value and average monthly value. Both the total and average values will be used to identify situations where the total value may be very close to a total value for an aggregate in the past, but the number of records that account for this value may be dramatically different. For example, if the total dollar value of imports from Germany contained in Chapter 14 was \$1,000,000 and was accounted for by 25 records, it would be suspicious if the closest value in the last twelve months was \$950,000 but was accounted for by only 3 records. Conversely, if 25 records from Germany for Chapter 14 had a total value of only \$5,000, this aggregate would probably be determined to be suspicious.

Both types of aggregates (Total Value and Average Value) will be used in the macro edit. For each type the value will be compared to the closest historical value within the last twelve months. A ratio of the current value over the historical value will be calculated. The ranking of the suspicious aggregate ratios will be based on the Hidiroglou-Berthelot method (1986).

There will be four groups with total value and average value aggregates associated with each group:

Group 1

1. HS10/Value
2. HS10/Average Value

Group 2

1. HS2/Importer/Value
2. HS2/Importer/Average Value

Group 3

1. HS2/Country/Value
2. HS2/Country/Average Value

Group 4

1. HS2/MOT/Value
2. HS2/MOT/Average Value

This method will transform the current value over the historical value ratio, to give equal weight to n-fold increases and decreases. This transformed ratio will then be weighted by a function of the maximum of either the current value or the historical value.

The fifty highest ranked ratios for each category of the four groups will be designated as suspicious. The next step will be to select aggregations that contain records for review. This will be accomplished by selecting a total of 200 aggregates, fifty from each group. For each group those aggregations that are ranked in the top fifty in both the total value and average value aggregations will be selected. If this results in less than fifty aggregates selected the remainder will be divided equally between the two type of aggregates.

For example, consider Group 1, which has thirty aggregations that were in the top 50 ranking for both the total value and the average value. These aggregates will be selected first. This leaves 20 aggregates; 10 of the next highest ranking will be selected from the value type and the remaining 10 from the next highest ranking average value type. Each group will follow this procedure, resulting in 200 aggregates being selected. From these aggregations 10 records (highest valued) will be identified for review.

There are two other situations of importance that have not been dealt with using the previous procedure. Aggregates that have a history, but are not present in the current month data, and aggregates that are present for the first time during the current month, with no previous history. A report of these aggregation will be produced, although they will not be part of a system aggregate/record review selection procedure.

The distribution of the records for review among the different attribute groups will be accomplished using a parameter table that allows for changes to the total number of records for review, the distribution to each of the four groups and an allocation among the number of aggregates in a group and the number of records selected from each aggregate.

Total Allocation = 2,000 (N)			
	Group Distribution	G r o u p Aggregates	Records in Aggregate
HS/CTY	GD = .40 times N = 800	40 (G)	RA = GD/G = 20
HS/IMPORTER	GD = .05 times N = 100	20 (G)	RA = GD/G = 5
HS/MOT	GD = .30 times N = 600	50 (G)	RA = GD/G = 12
HS2	GD = .25 times N = 500	100 (G)	RA = GD/G = 5

The preceding table allows for an adjustment to the distribution of the records by macro group, and an adjustment to the number of aggregates within each group or the number of records within each aggregate or both.

6.7 Imputation module

All micro failures identified will be corrected during the review process. All records that had not been selected for review from the selective, pre-macro (objective), pre-macro (subjective) and the macro edit modules and have micro level failures will be resolved by system imputation. The imputation is deterministic and will use a combination of metadata tables and information from the field values to determine a value for the micro edit failure.

6.8 Post imputation macro edit module

After the completion of the imputation module, the data will be subjected to the same macro edit conditions. The

objective of the post imputation macro edit is to verify that the changes to the data for records that were reviewed and changes resulting from imputation have not created new suspicious aggregates. Review of records will only occur if extreme changes in aggregate values have resulted from the process.

6.9 USBC edit module

The International Trade Division of Statistics Canada, has a Memorandum of Understanding with the Foreign Trade Division of the United States Bureau of the Census, whereby each statistical agency supplies the other with their imports data for that country, which become the exports statistics for the receiving country. With this agreement come requirements from both International Trade Division and the Foreign Trade Division for certain editing to take place in the other's import processing. The USBC editing module exists to fulfill requirements that the Foreign Trade Division has for a few select fields in Canadian imports data. This will be the last module in the Edit and Imputation System because other fields that affect these edits and determine if they will be performed must be validated, reviewed, and corrected and any edit failures result in deterministic imputation. There are only four fields that will be edited in this module: Port of Unlading, United States Port of Exit, Weight and Freight.

7. Review procedure

The records that have been identified for review from the different edit modules will be presented to the data quality analysis staff for resolution. Review records will be allocated using the chapter information to one of nine specialty areas. For example, Specialty 1 which is made up of Chapters 1-24 and 91-98.

The review modules contain information on the number of records for review and the total value of these records and will be arranged by attribute (aggregation) where appropriate. Any records that have been identified for review in the exclusion module (executed before the edit and imputation process) will be distributed using the chapter information to the corresponding specialty area. The selective review will be presented in the same manner as the exclusion, assignment to specialty area.

The pre-macro (objective) will, in addition to using the chapter information to assign the review records to specialty areas, further group the records by aggregates. Thus all records that have been identified for review relating to one of the seven micro edit failure will be grouped together. Each specialty area will have the review records grouped by these attributes.

The attribute groups used in the pre-macro (subjective) will be the individual rules used to monitor the imports data. After the review records have been assigned to their specialty area, they will be grouped by monitoring rules. Similarly the macro review in the specialty areas will be grouped by the aggregations used in the macro edit; HS2/country, HS2/mode of transport, HS2/importer and HS10.

The document retrieval function for the imports system assumes that not all lines requiring review will need the paper documentation to adequately perform that review and make the necessary corrections. For example, if the "error" being reviewed is one that occurs regularly, the subject matter officer may already know where the problem is and can make the correction without referring to the documents. The process allows the review officer to determine which specific lines can only be verified with the paper documentation or electronic invoice information. This will reduce the workload of manually pulling documents while not affecting the quality of the review. In addition to a smaller number of documents being requested, these requests will most likely be staggered throughout the month, rather than large requests once or twice a month.

8. Disadvantages and constraints

Although the development and the testing of a new system is initially expensive, it pays for itself in the long run considering all its advantages previously discussed. But one question remains: how this overall system will perform? For the moment,

that question has only been partially answered. The results of the selective edit module simulations were very encouraging and we hope that those from the subsequent modules will go in the same way. The available prototype of the system will allow us to evaluate each module individually and also together to find the best one.

Finally, the staff training will probably be one of the most important challenges of the new E&I process. It will be essential that the staff associated with the edit and imputation system be comfortable enough with the new approach to promote and build over time, a continually improving system applicable to the constantly changing needs of the ITD program. It is why we have to make sure that all necessary efforts will be done on the training of the staff to ensure their skill development and ability to improve the efficiency of the system.

9. Conclusion

The re-engineering of the edit and imputation system will provide ITD with a very different approach to verifying the consistency, integrity and quality of the data with the joint use of micro and macro approaches in the same editing system. The development of the new strategy has been completed, the design and delivery of the modules is underway with a planned delivery for User Acceptance Testing for the summer of 1996. These modules will be evaluated individually, integrated and then subjected to a system test. It is known that the new process will provide a more focused and targeted approach to records that have significant impact on the trade statistics. Furthermore, the modularity of the system allows for more flexibility in incorporating improvements into the new system.

The complementary review system presents the contributors belonging to suspicious aggregates to be viewed as a complete group. Because of the group review of records as opposed to the line by line review and the ability of the new system to more effectively target suspicious records, a savings in the amount of time required to review data has been identified.

Anticipated gains are envisioned due to increased use of electronic information for the automotive parts records and Customs' electronic invoices. Both of these files will be used in the review of suspicious records and can be accessed by 'clicking' on a indicator. The instantaneous retrieval of the parts and electronic invoice relating to a transaction during the review process, results in savings of both time and paper.

The event and audit logs which record all edit failures and any changes to any of the data fields contain the information required to perform a complete analysis of the impact of the edit and imputation system.

In many areas, the new edit, imputation and review processes will mean considerable changes to the scope of individual responsibilities within ITD. A successful transition to the new organization required for this new approach is crucial to its implementation.

In summary, the reduction in the number of records identified for review, the decline in the number of paper documents requiring retrieval and the improved ability to monitor the impact of the processing, all translate into process optimization and subsequent resource savings.

Bibliography

- Chinnappa, N., R. Collins, J.-F. Gosselin, T. S. Murray, and C. Simard (1990), "*A Report of the Working Group on Strategies for Macro editing*," unpublished report, Ottawa, Statistique Canada.
- Granquist, L., and Kovar, J. G. (1994), "*Edit of Survey Data: How much is enough?*", Paper submitted for the Monograph of International Conference on survey Measurement and Process Quality, Bristol, UK, April 1995.
- Hidioglou, M.A., and J.-M. Berthelot (1986), "*Statistical Edit and Imputation for Periodic Surveys*," Survey Methodology, 12, pp. 73-83.
- Laflamme, F. (1995), "*La Macro-vérification dans le cadre du programme du Commerce International*," Proceedings of the ACFAS conference, Chicoutimi, May 1995.
- Laflamme, F., Barrett, C., Johnson, W. and Ramsay, L. (1996), "*Expériences dans la restructuration de la stratégie de vérification et d'imputation des importations canadiennes*" Proceedings of the ASU conference, Québec, May 1996 (to be published).
- Laflamme, F., Barrett, C., Johnson, W. and Ramsay, L. (1995), "*New Import and Imputation System in ITD*," unpublished report, Ottawa, Statistique Canada.
- Pursey, S. (1994), "*Current and Future Approaches to Editing Canadian Trade Import Data*," Proceedings of the Survey Research Methods Section, American Statistical Association